Leveraging computer vision systems for monitoring animal health and productivity on dairy farms

ARIANA NEGREIRO
PHD STUDENT
ADVISOR: DR. JOAO DOREA

DEPARTMENT OF ANIMAL & DAIRY SCIENCES
UNIVERSITY OF WISCONSIN-MADISON

Department of
Animal & Dairy Sciences
UNIVERSITY OF WISCONSIN-MADISON

My Background

Animal scientist

Focus on animal health and management

CHTC allowed me to accelerate my research and perform analyses that would be impossible otherwise

Outline

Research — My research

Implementation — Computational implementation

- computing requirements
- deployment
- throughput/time

CHTC — Using CHTC/HTCondor

Impact — Personal and professional impact

Department of
Animal & Dairy Sciences
UNIVERSITY OF WISCONSIN-MADISON

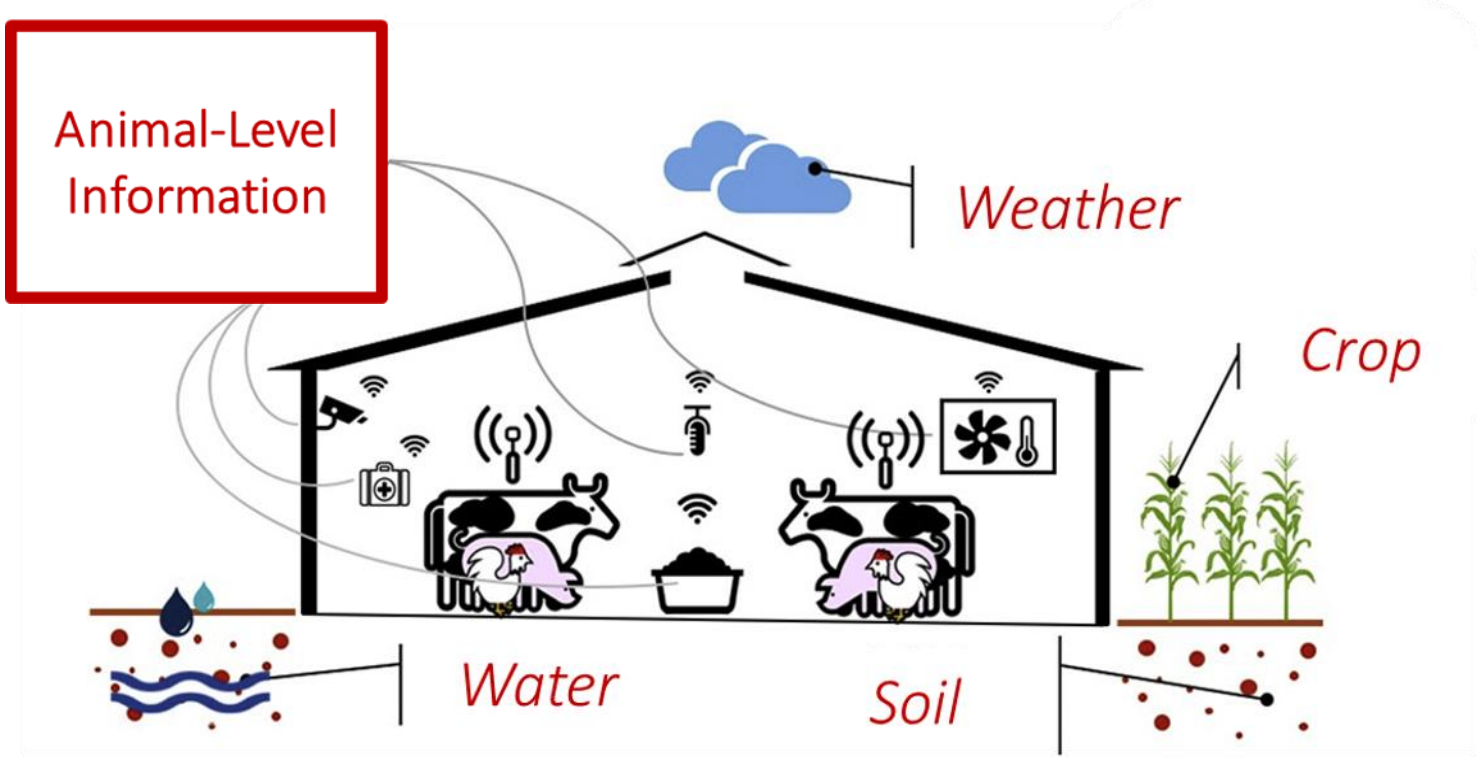
Dr. Dorea's Lab

Research | Implementation | CHTC | Impact

- Research applications of machine learning and computer vision for farm management and genetic selection

Sensors:
Wearable
Cameras
IR Spec.
RFID
Sound
Housing

Animal Identification
Animal Behavior
Body Weight
BCS/Composition
Milk Components
Milk Yield
Estrus Event
Feed Intake
Feed Efficiency
Disease Risks



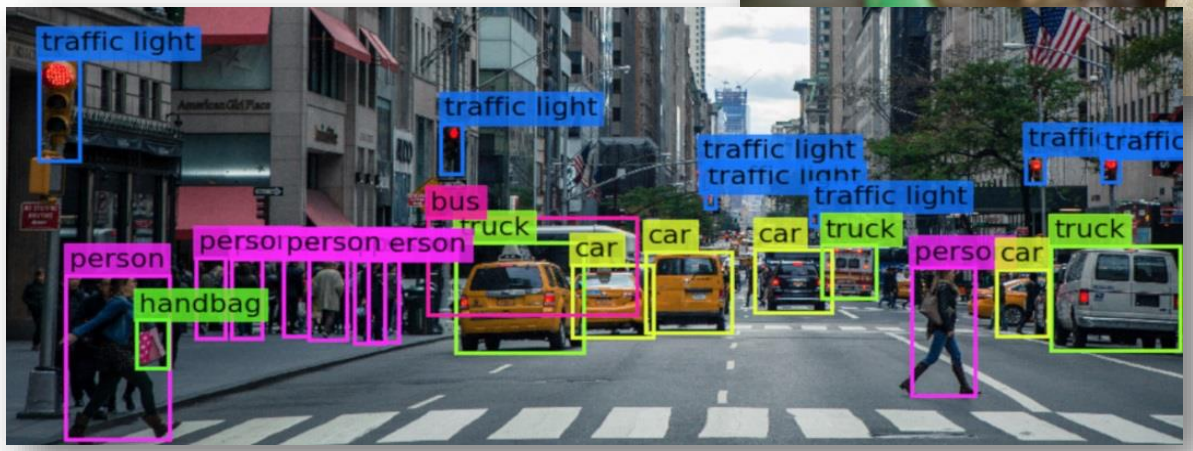
Tullo et al., 2019

Monitoring Behavior in Livestock

Current methods for monitoring cattle:

- Visual observation
  - Large-scale applications?
  - Subjectivity
- Wearable sensors

* Computer vision



Benefits of Computer Vision

- Cameras are affordable and easy to install
- Ability to monitor multiple animals at a time
- Images provide a great amount of information
  - Animal location/action
  - Social interaction
  - Weather/season
  - Health/mobility

Department of
Animal & Dairy Sciences
UNIVERSITY OF WISCONSIN-MADISON

Research: Cow mouth tracking

Research → Implementation → CHTC → Impact

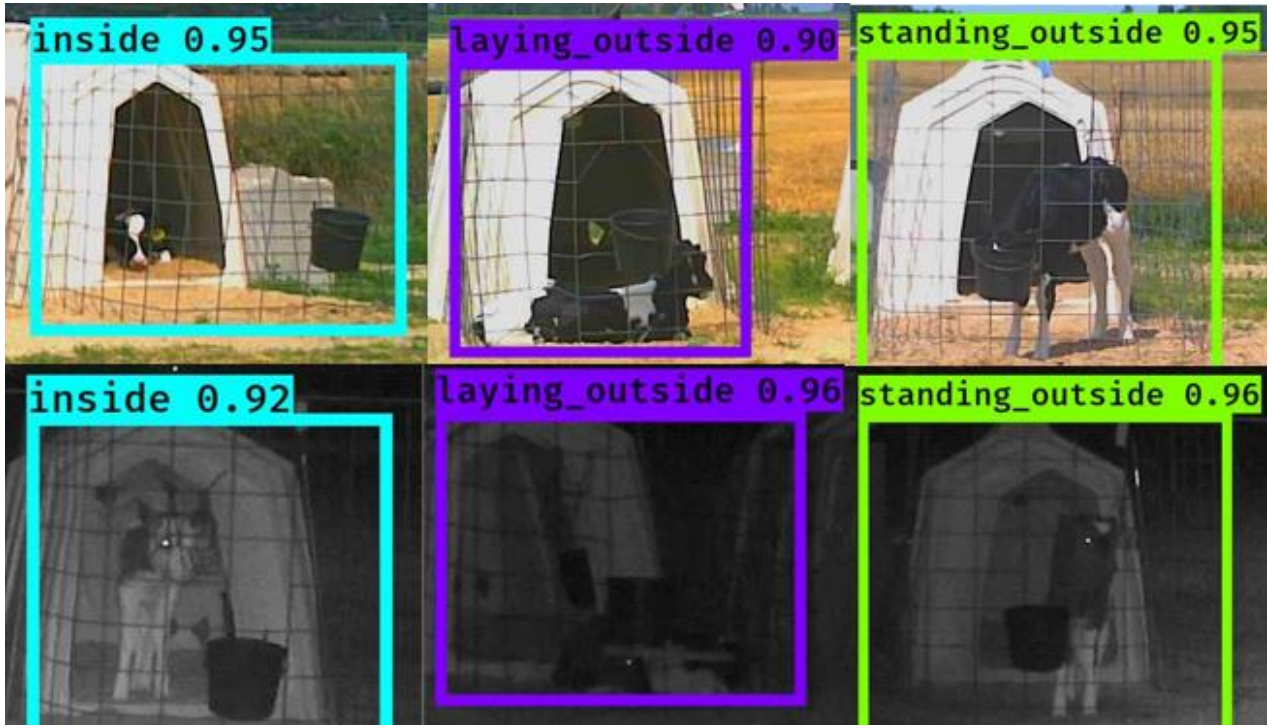
- 1,662,417 images
  - Cropped for each individual cow (utilizing CHTC)
    - 4,008,630 sequential images
    - Mask R-CNN

Training
348 images

Validation
37 images

Inferences made on remaining 4,008,245 images

Research: Calf heat stress detection



- 27,704 images
- YOLOv3
- tinyYOLOv3

Training
297 images

Validation
128 images

Inferences made on remaining 27,704 images

Research: Calf identification and growth monitoring

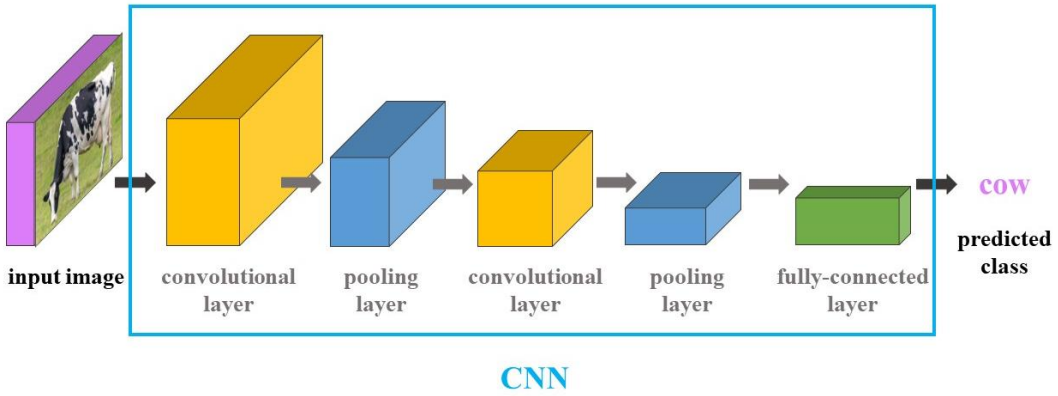
Research → Implementation → CHTC → Impact

- 27 + TB of data
- Current project
- Xception (ongoing analyses)

Department of
Animal & Dairy Sciences
UNIVERSITY OF WISCONSIN-MADISON

Implementation

- These analyses utilize complex algorithms



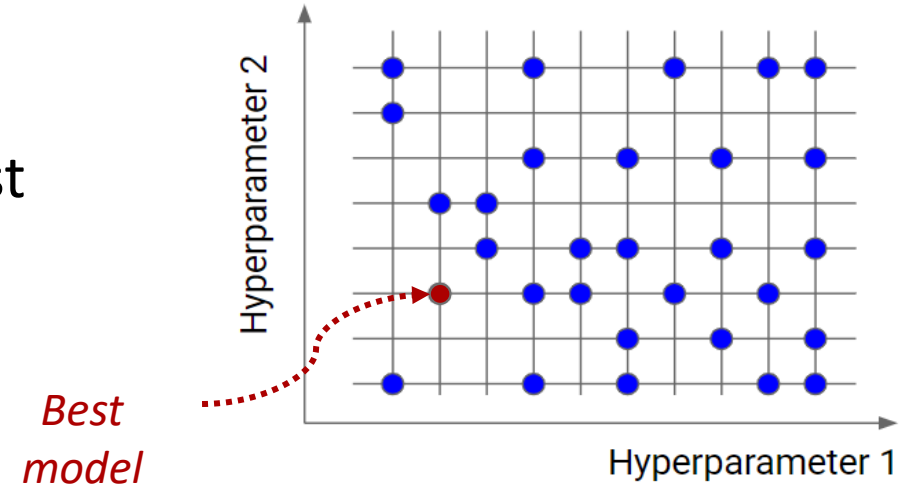
- Large datasets (thousands or millions of images)
- Image data is large (400x600 = 24,000 pixels)
- Image preprocessing
- Computationally demanding to train

- Mask R-CNN
  - 44 million parameters
- YOLOv3
  - 40.5 million parameters
- tinyYOLOv3
  - 8.9 million parameters
- Xception
  - 22.8 million parameters

Neural Network Training

- Multiple neural networks to train independently:

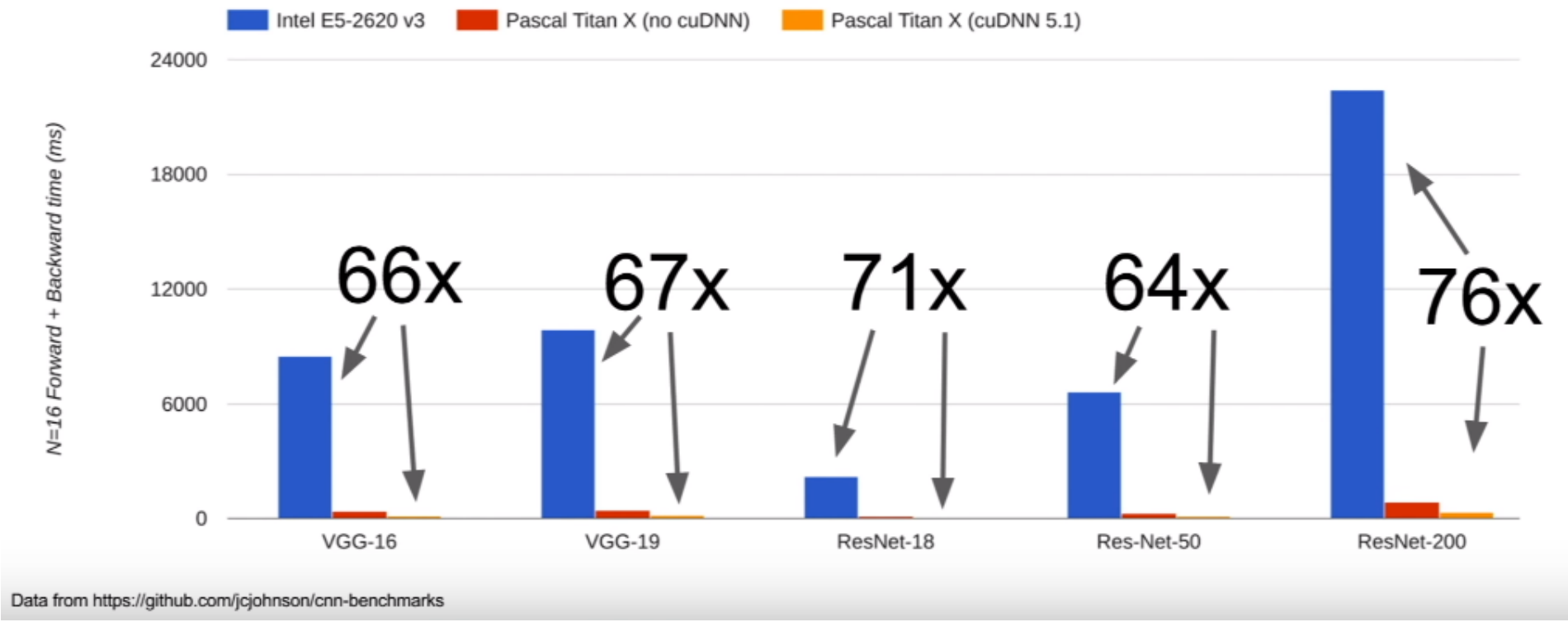
- Different datasets
  - Evaluate which one is the best for training (preprocessing, data collection strategies, etc)
  - Perform multiple experiments (effect of day, lighting, etc)
- Hyperparameter tuning
  - Train using multiple combinations to find the best



Neural network training: GPUs vs CPUs

CPU vs GPU in practice

(CPU performance not well-optimized, a little unfair)



Stanford cs231n.

Using CHTC/HTCondor

- Using since Fall 2021
- > 4,000 computing hours
- Larger projects require dozens of trained neural networks
- Each project contains thousands of images for training, and 100,000s or even millions of images for inference

Using CHTC/HTCondor

- Datasets stored on Staging
- Python environments using Miniconda
  - Compressed environment packs stored on SQUID
- Queue jobs using txt files

```
knn_iteration0_60_120_2048_000,xception,30,60,8,59
knn_iteration0_60_120_2048_050,xception,30,60,8,59
knn_iteration0_60_120_2048_075,xception,30,60,8,59
knn_iteration0_60_120_2048_100,xception,30,60,8,59
xgboost_iteration0_60_120_2048_000,xception,30,60,8,59
xgboost_iteration0_60_120_2048_050,xception,30,60,8,59
xgboost_iteration0_60_120_2048_075,xception,30,60,8,59
xgboost_iteration0_60_120_2048_090,xception,30,60,8,59
```

- Template folder for each category of project (containing .sub, .sh, python files, etc)
  - Each template expects datasets following a certain format and outputs files/folders following a certain format

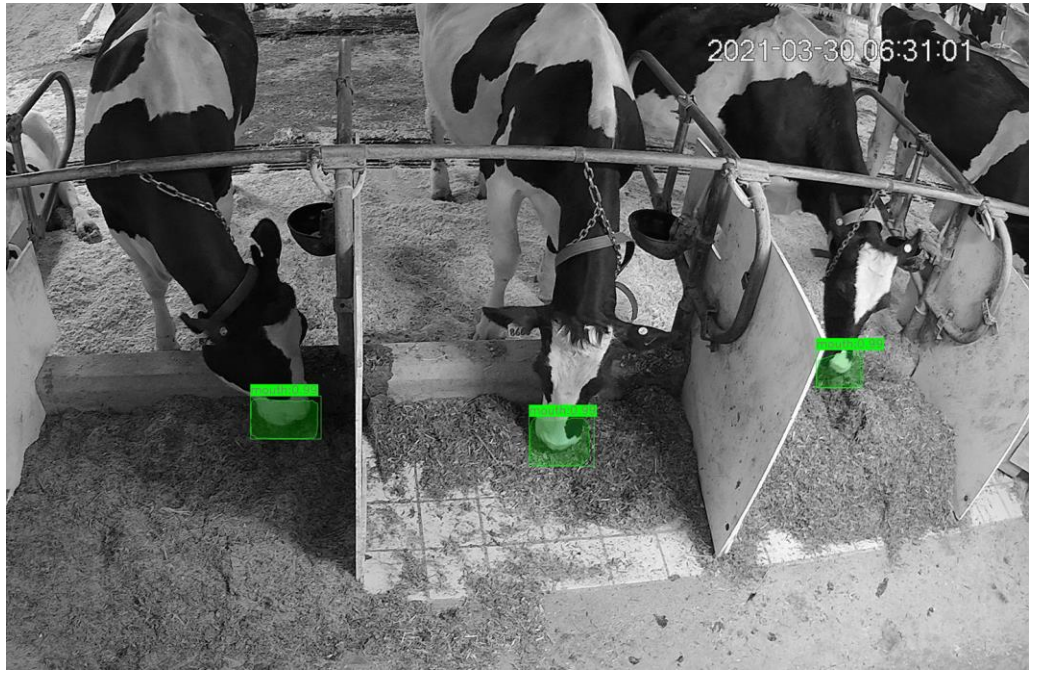
Department of
Animal & Dairy Sciences
UNIVERSITY OF WISCONSIN-MADISON

Use case: cow mouth detection

Research → Implementation → CHTC → Impact

Input: Dataset containing images and bounding boxes

```
Dataset/
    Img01.png
    Img01_bbox.txt
    Img02.png
    Img02_bbox.txt
    Img03.png
    Img03_bbox.txt
    ⋮
```



physical measurements
- distance traveled
- acceleration
- velocity

provides insight on feeding behavior, health, and milk production

- Output: Bounding box predictions on test set

Department of
Animal & Dairy Sciences
UNIVERSITY OF WISCONSIN-MADISON

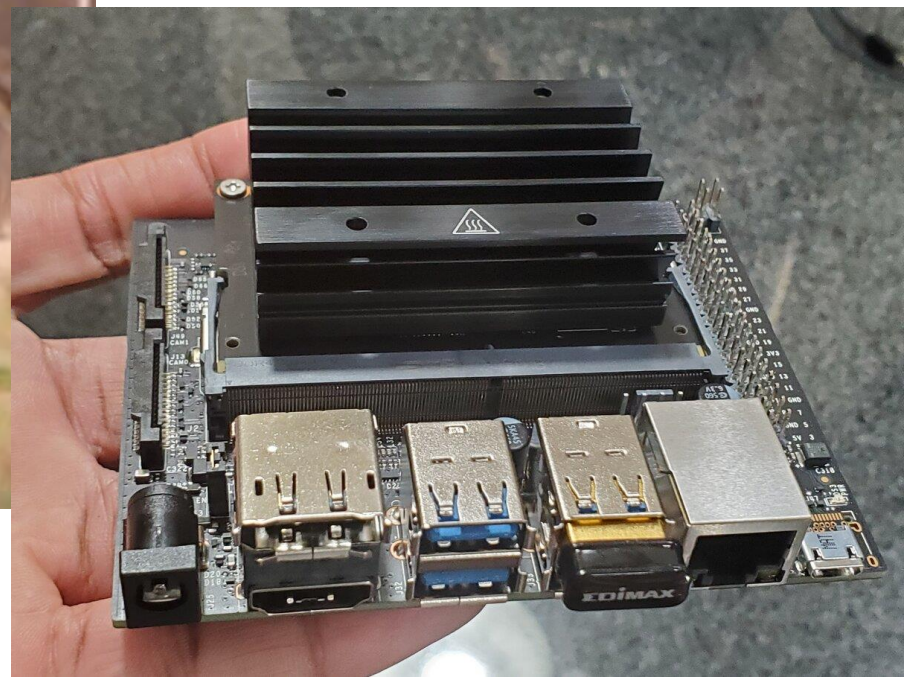
Use case: calf detection

Research → Implementation → CHTC → Impact

- Input: Dataset containing images and masks
- Output: Mask predictions on test set

Use case: edge computing

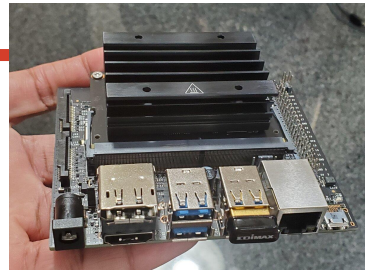
Research → Implementation → CHTC → Impact



Deploy trained model (trained using CHTC) in edge computing applications to make inferences real-time

Use case: calf identification

Research → Implementation → CHTC → Impact



- Collecting images 24/7 following calves the first 2 years of life
- Tracking growth, health, and behavior
- Estimated > 100 TB data to be acquired

Department of
Animal & Dairy Sciences
UNIVERSITY OF WISCONSIN-MADISON

CHTC Benefits/Limitations

Pros:

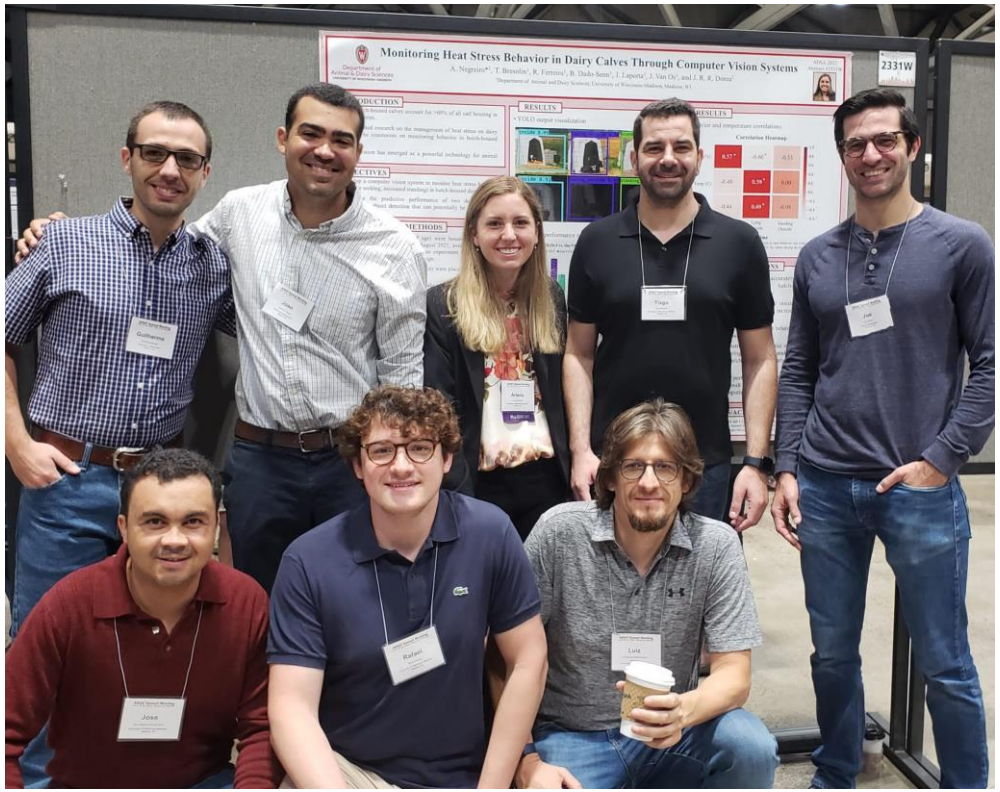
- Potential ability to access data directly from our own servers (access point)
- Checking logs to have an idea of how far into the job (which epoch, for example) the 12/24/72hr limits were reached
- Flexibility to submit jobs to CPUs or GPUs depending on availability and size of job
- The option to have emails sent when jobs are done running

Cons:

- Large datasets can take very long to transfer, especially when working from home using a VPN

Personal & Professional Impact

- Ability to perform data analysis that would be impossible otherwise
  - Advancing knowledge of animal health and behavior
- Experience accessing a remote Linux server
- Exposure using a high-throughput computing system
- Consider data flow and automation within remote server environments
- Cultivated skills that will help me in my future career, opened opportunities to present my research in multiple conferences, and work with a great team!



Thank you!

negreiro@wisc.edu

Department of
Animal & Dairy Sciences
UNIVERSITY OF WISCONSIN–MADISON